5-1-1983

An application of multi-dimensional scaling to the re-examination of preferred tone reproduction

David J. Porter

AN APPLICATION OF MULTI-DIMENSIONAL SCALING TO THE RE-EXAMINATION OF PREFERRED TONE REPRODUCTION

by

David J. Porter

B. S. North Carolina State University at Raleigh

(1972)

A thesis submitted in partial fulfillment
of the requirements for the degree of
Master of Science in the School of
Photographic Arts and Sciences in the
College of Graphic Arts and Photography
of the Rochester Institute of Technology

May, 1983

Signature of the Author......David J. Porter......
Photographic Science
and Instrumentation

Accepted by......Signature not legible......
Coordinator, Graduate Program

School of Photographic Arts and Sciences
Rochester Institute of Technology
Rochester, New York

CERTIFICATE OF APPROVAL

MASTER'S THESIS

The Master's Thesis of David J. Porter
has been examined and approved
by the thesis committee as satisfactory
for the thesis requirement for the
Master of Science degree

....
Doctor Edward M. Granger, Thesis Advisor

....
Professor John Carson

....
Doctor Kenneth Wommack

7/14/83
....
Date

THESIS RELEASE PERMISSION FORM

ROCHESTER INSTITUTE OF TECHNOLOGY
COLLEGE OF GRAPHIC ARTS AND PHOTOGRAPHY

Title of Thesis: An Application of Multi-dimensional Scaling to the Re-examination of Preferred Tone Reproduction

I, David J. Porter, hereby grant permission to the Wallace Memorial Library of R. I. T. to reproduce my thesis in whole or in part. Any reproduction will not be for commercial use or profit.

David J. Porter 5/31/83

AN APPLICATION OF MULTI-DIMENSIONAL SCALING TO THE RE-EXAMINATION OF PREFERRED TONE REPRODUCTION

by

David J. Porter

Submitted to the Photographic Science and Instrumentation division in partial fulfillment of the requirements for the Master of Science degree at the Rochester Institute of Technology

ABSTRACT

Black and white prints were produced of several scenes that varied in both negative contrast and negative exposure. Subjects rated the various prints by categories based on their own personal understanding of what an excellent print was. The boundaries of the categories were found and used to determine the preferred tone reproduction characteristics. Individual differences were also scaled with the use of a proximity measure determined from the categorical data. The preferred contrast (1.46 $\pm$0.24) was significantly higher than the contrast that has been found to be preferred in earlier studies. The application of this study to more general situations is also discussed.

ACKNOWLEDGEMENTS

Dr. Edward M. Granger of the Eastman Kodak Company provided me with the assistance and advice necessary to carry out and to complete this project.

The Central Intelligence Agency provided funds for the materials used in the production of the many prints needed for this study.

Richard Norman of Rochester Institute of Technology provided miscellaneous mechanical parts and advice.

Professor John Carson and Doctor Ron Francis of Rochester Institute of Technology provided several ideas central to the execution of various portions of the experiment.

My parents have supported my decision to embark on the course that has lead to this project and have provided both monetary and psychological support along the way.

My wife, Kathleen, has patiently put up with this project and its demands since its inception.

G9289986

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

INTRODUCTION

Tone reproduction is concerned with the reproduction by photographic means of the appearance of an object, assuming that image structure is not a factor. Objective tone reproduction refers to the photographic reproduction, point by point, of scene luminances and subjective tone reproduction refers to the photographic reproduction, point by point of brightness, that is, the psychophysical response to luminance.

ORIGINS

The origins of tone reproduction date back to the beginnings of photography. M. Hubert, Daguerre's assistant, published an exposure guide in 1839 and D.W. Seager, the first person to make a daguerreotype in America, published a more complete exposure guide in 1840.[1,2] These guides helped early Daguerreotypists achieve a reasonable exposure on their plates and, therefore, a reasonable tone reproduction. As the wet plate processes came on the scene, exposure guides were still useful because the sensitivity of most emulsions were fairly constant even though the photographer himself produced the sensitive surface.

The dry plate era, however, produced photographic emulsions that varied widely in speed. This was due primarily to variations in the little understood process of ripening and small impurities in the gelatin that increased the spectral sensitization and therefore the speed of the primitive silver halide. F. Hurter and V. C. Driffield spent several years studying the photographic process and formulated a relationship between the chemical density of silver per unit area and the exposure incident on that area.[3]

For photographic purposes, density has come to mean optical density, or the light stopping power of the silver. It is measured both in the print and on the negative by reflection and transmission methods, respectively. Instruments for measuring density and for exposing film have become very precise, providing the careful worker with easily obtained curves relating the input, exposure, to the output, density, that is characteristic of the material being tested.

TONE REPRODUCTION CURVES

If the characteristic curves for all steps in a photographic process are available, these curves can be cascaded, producing a curve that relates the system input to the system output. This curve is known as the tone reproduction curve and is said to be a subjective tone reproduction curve if the input and output are in brightness units; objective, if the input and output are in luminance units.

The luminance of the print, for any given reflection density, is proportional to the level of incident illumination. To avoid having print illumination be an important contribution to each objective tone reproduction curve, the reflection density is often used as the output of the system rather than print luminance.

IDEAL TONE REPRODUCTION

It was first thought that the ideal tone reproduction would exactly reproduce the luminance of the object. This corresponds to an objective tone reproduction curve that is a 45° straight line (Figure 1).[4] The print and scene would

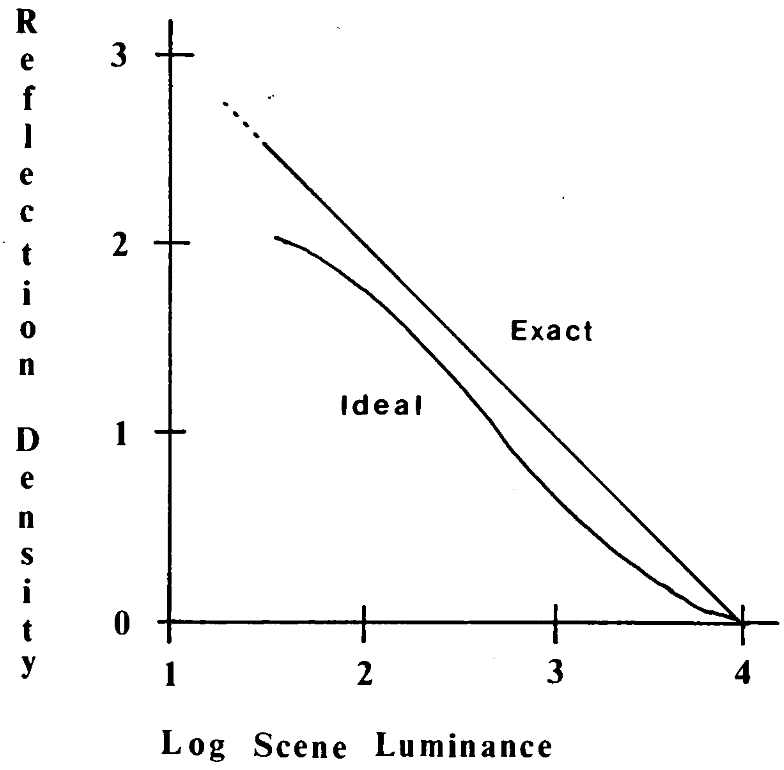


FIGURE 1. IDEAL TONE REPRODUCTION

Ideal tone reproduction is shown to be lighter (0.3) and to have a higher gradient (1.10) than exact tone reproduction.

have to be illuminated identically. Prints are seldom viewed this way, so most of the experimental work has been done in the more normal illumination of 100 footcandles. The experiments of Jones and Nelson in 1942[5] and Jones and Condit in 1948[6,7] for prints viewed normally, showed a preferred objective tone reproduction curve that was lower in density than the supposed ideal 45° line by nearly 0.3 and had a midpoint gradient of 1.10 rather than 1.00 (Figure 1).[4]

It was found by C. D. Edgett, K. C. Whener, and C. N. Nelson that the density discrepancy was due to the inability of the eye to compensate for the absolute difference in luminance between the scene and the print when the luminances of the print were 1/100 of the luminances of the scene.[9] The eye could compensate when the luminances of the print were 1/50 of the luminances of the scene. Therefore, a print that was made lighter than exact luminance reproduction would allow would be favored in viewing situations with lower illuminations. A higher midtone reproduction gradient is necessary in this case to give a better match to the ideal exact luminance reproduction curve.

VISUAL RESPONSE

Obviously, the response of the eye to luminance is a large contributor to tone reproduction quality. G. T. Fechner in 1889 concluded that the eye responds logarithmically to luminance.[10] R. B. Marimont has shown that brightness more closely follows a power law equation of the form:[11]

$$B = k\ (L/L_{av})^{n} \qquad (1)$$

Where:

B = Brightness

k = Gain factor based on surrounding luminance

L/L_{av} = Simultaneous Contrast

n = Power

This equation takes into consideration both adaptation of the eye to the general surrounding luminance and the simultaneous contrast effects where the difference in luminance between a dark area and a surrounding light area produce a larger brightness difference. This model accounts for much of the effects of different surrounds of photographs and the heightened contrast or Mach effect noticed at the edges of different patches of uniform density. Because the uncertainties in calculating the constants of the equation are so large, an absolute subjective model cannot be formulated and accurate subjective tone reproduction curves based on this model are not possible. Despite this, the model does give a good qualitative insight into the response function of the eye.

RECENT TONE REPRODUCTION WORK

Tone reproduction curves are laborious to produce. If an experiment uses quality as a response, many prints are necessary to judge quality psychophysically, each requiring a tone reproduction curve. J. Simonds[12] developed a computer program that derived the tone reproduction curves digitally. In addition, psychophysical quality ratings were regressed onto the "characteristic vectors" of each tone reproduction curve. By using the regression equation, an experiment could be simulated that would have otherwise required the creation of many carefully controlled prints and voluminous data from many judges.

PROBLEM

Simonds' method allows for the optimization of each significant factor concerning tone reproduction, but only a limited amount of data from his work is available. His method of calibration involved the use of quality data that was regressed onto three vectors that were characteristic of the tone reproduction curve. Clark's later work, is philosophically identical except that monochrome transparancies were used to allow an effective reflection density greater than 2.0 and that the regression included four characteristic vectors: the three mentioned above and one for screen luminance.[13]

Nelson had judges view prints for the purpose of finding an optimum safety factor for camera exposure.[14] In this study, quantitative judgement data was used to provide a unidimensional scale of quality directly. Exposure was optimized, but contrast was not a factor.

Jones and Nelson's work evaluated photographic printing relative to negative characteristics.[5] Quality was scaled by presenting judges with a field of fifteen prints ordered such that contrast increased by row from top to bottom and exposure increased by column from left to right. Judges picked the best print in each row and then rank ordered these prints. The contrast grade yielding the highest quality was found by weighting the frequency that each contrast grade was chosen according to the rank order. Likewise, the best exposure was the weighted average of exposures on the best contrast grade. This scaling procedure nested contrast within exposure and forced the three best prints into mutually exclusive contrast grades. The resulting quality scale was correlated with several characteristics, e.g. density scale vs. contrast grade and exposure vs. minimum density.

In all the studies mentioned above, quality was considered a unidimensional quantity. Variability in quality was or forced onto one or more given parameters. The nesting of parameters and forcing the best choices into mutually exclusive contrast grades may produce biased results. There is some evidence that newer photographic systems such as instant color photography have midpoint tone reproductions significantly larger than the 1.15 gradient that Jones and Nelson found was optimum in 1942.[9] This suggests that an experiment to re-examine this area may produce some additional knowledge.

THESIS

My thesis is that the preferred exposure and midpoint gradient of the tone reproduction curve for a general landscape scene may be different from those found to be preferred previously. The hypothesis that will be tested is that the optimum exposure and midpoint tone reproduction curve gradient is the same as previously found.

SCOPE

My interest in this project grew from my early failures at obtaining high quality black and white photographs of landscape scenes. This work was done in the 4x5 format and I wanted to use this format in the hope of applying my findings to my own work. Many other personal biases have affected the design of the experiment in several ways. One of these is the aversion I had to the inclusion of psychological keys in the scenes such as an object that would be assumed to be white or skin tone. However, more subtle keys were included in the scenes. These include grass, sky, trees, concrete, water, and stones. While the actual appearance of these objects can change drastically with differing lighting conditions, they do provide the subject with subtle keys of what the picture should look like.

Another personal bias I had was to perform the tone reproduction control on the negative only. This has its origin in the Zone System as described by Ansel Adams, Minor White, and others.[15,16] In this system, the processing and exposure of the negative is adjusted for the contrast and the absolute luminance of the scene, respectively. Ideally this procedure will produce a negative that requires very nearly the same printing exposure and contrast grade paper as a similarly produced negatives of any other scene.

Because an opal glass densitometer reads an optical density that is an approximation to integrating sphere density and not the density that a printing system would see, it was considered desirable to use contact prints in this study.[17] The 4x5 format was large enough to allow this and it also simulates rather well the 3 1/2 x 5 inch prints that are currently popular.

This simulation was used to advantage in the actual viewing of the prints. Initially, I had wanted to control very carefully the viewing conditions in both geometry and level of illumination. However, prints are seldom viewed that way. In general, prints are viewed in the ambient illumination with only small consideration being taken of geometry (eliminating glare by proper orientation of the print). In this study the viewing conditions were monitored, not controlled and did vary widely.

EXPERIMENT

OVERVIEW

This experiment attempted to measure the psychophysical quantity of print quality and to determine from those measurements the objective tone reproduction that was most preferred. Prints were produced as stimuli that varied systematically in log exposure from -0.2 to +0.2 in 5 levels centered on the recommended film exposure and in contrast from 0.62 to 1.49 also in 5 levels. A single film-developer combination was used with contrast controlled by varying the concentration of developer and the time of development. Figure 2 graphically shows the effect on the tone reproduction curve of the extremes of both contrast and exposure. The stimuli were presented to subjects for judging on a "quality" scale. This quality scale was not impressed on the subjects, but was psychophysically measured by letting each subject choose the quality based on his own judgement. No subject coaching or limiting of what the subject was to consider took place. Because 5 levels of exposure and 5 levels of contrast were used, paired comparisons of each pair of stimuli for each

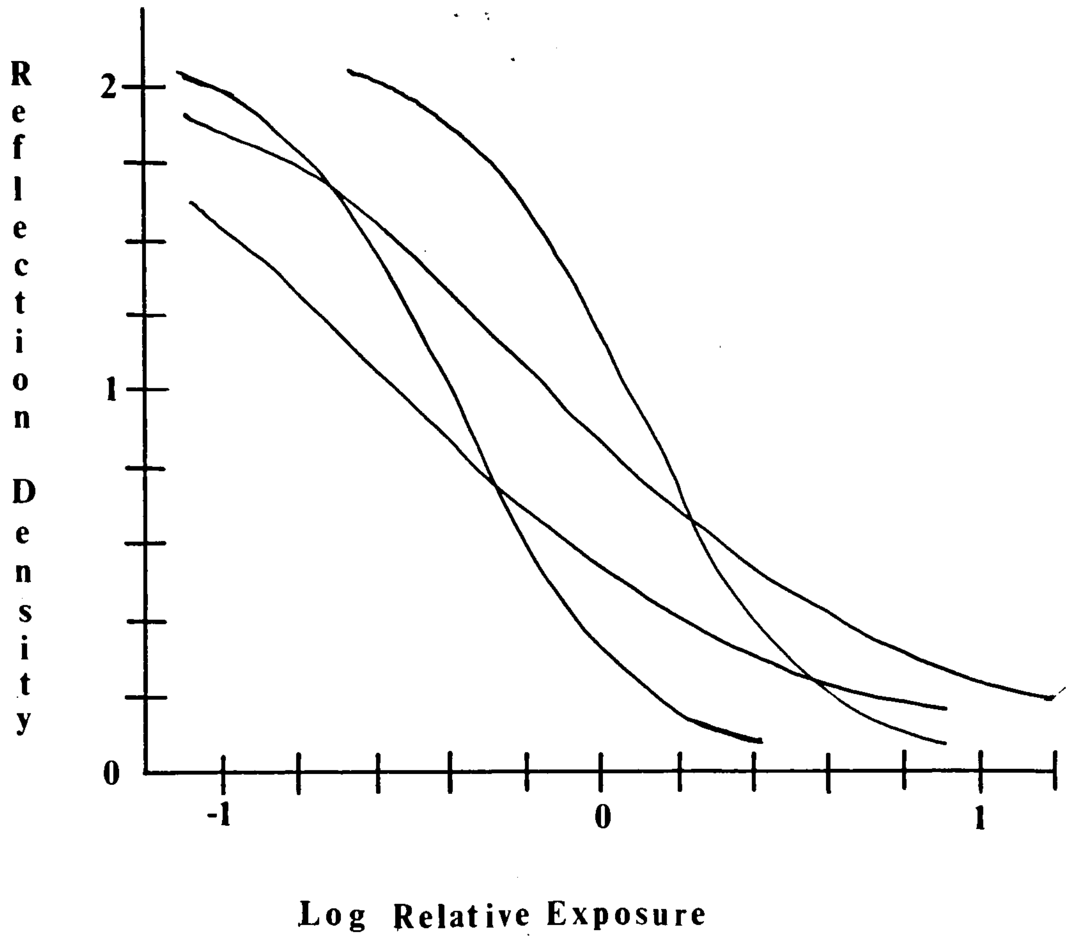


Figure 2

TONE REPRODUCTION CURVES FOR EXPOSURE AND CONTRAST EXTREMES of 3 scenes would have required far too much time and effort from the subjects. In order to speed this part of the experiment up, the law of categorical judgement was used . The subjects were told only to judge each print on a scale of 1 to 7. Some were given a list of adjectives that have proven useful in this type of experiment in the past[8], and some were not. The discriminal dispersions of the category boundaries on the psychological continuum were determined and contour plots made from the results. From these data, preferred tone reproduction characteristics were determined.

In addition to the above analysis of the judgement data, a multi-dimensional scaling of the individual differences between judges and individual differences between stimuli was carried out. Because the input to this type of analysis must be some measure of the proximity of two stimuli (similarity or dissimilarity) a measure that is monotonic to proximity was constructed. This was done by subtracting the category assigned by a subject to each print from the category assigned by the same subject to every other print for each scene. These data were used by an INDSCAL program for computation. The primary reason to carry out this analysis was to allow the detection of subjects that seemed to be judging in an atypical manner. Nevertheless, it was hoped that the results of this would provide new insight into the problem from the knowledge of the dimensionality of the response.

PRODUCTION OF STIMULI

The prints to be used as stimuli for this experiment had to be produced to very precise standards. An Arca-Swiss view camera was modified for the purpose of taking the negatives.

CAMERA SET-UP

The camera used for this experiment was a specially modified Arca-Swiss 4x5 view camera with a 150 mm f-5.6 Symmar lens. Because a view camera provides many degrees of freedom that were not needed, the camera was modified and aligned prior to the experiment. No further adjustments were made to the camera until all negatives had been made. The following items were adjusted: Aperture, Alignment, Focus, Lens shade, Film plane mask.

APERTURE

The choice of the camera system aperture is one of the most important decisions regarding image quality. The aperture controls depth of field, image sharpness, shutter speed, shutter efficiency, and spherical aberration to name only a few. Because of the many effects that the aperture can have on the system, there are some contradictory constraints that must be dealt with. A small aperture is dictated by the desire for a great depth of field, high shutter efficiency, and low spherical aberration. A large aperture is dictated by the desire to have as small a diffraction circle and as short a shutter speed as possible.

<u>EXPERIMENTAL TEST.</u> A test of the image sharpness at the full range of f-stops available was carried out to determine if images at infinity (the far limit of depth of field) would be visibly degraded. Three collimators with Siemans star targets were set on an optical bench and carefully autocollimated. Six exposures were made, one at each f-number. The negative was inspected and it was determined that f-45 would be usable for contact prints.

GEOMETRICAL OPTICS. Formulas from geometrical optics were used to determine whether the depth of field would be sufficient at f-45.[18]

$$HD = LA / AB \tag{2}$$

$$NL = FL^2 / (2\ f\#\ BC) \tag{3}$$

where:

HD = Hyperfocal Distance

LA = Linear Aperture

AB = Angular Blur

NL = Near Limit of Focus

FL = Focal Length of Lens

f# = f-Number of Lens

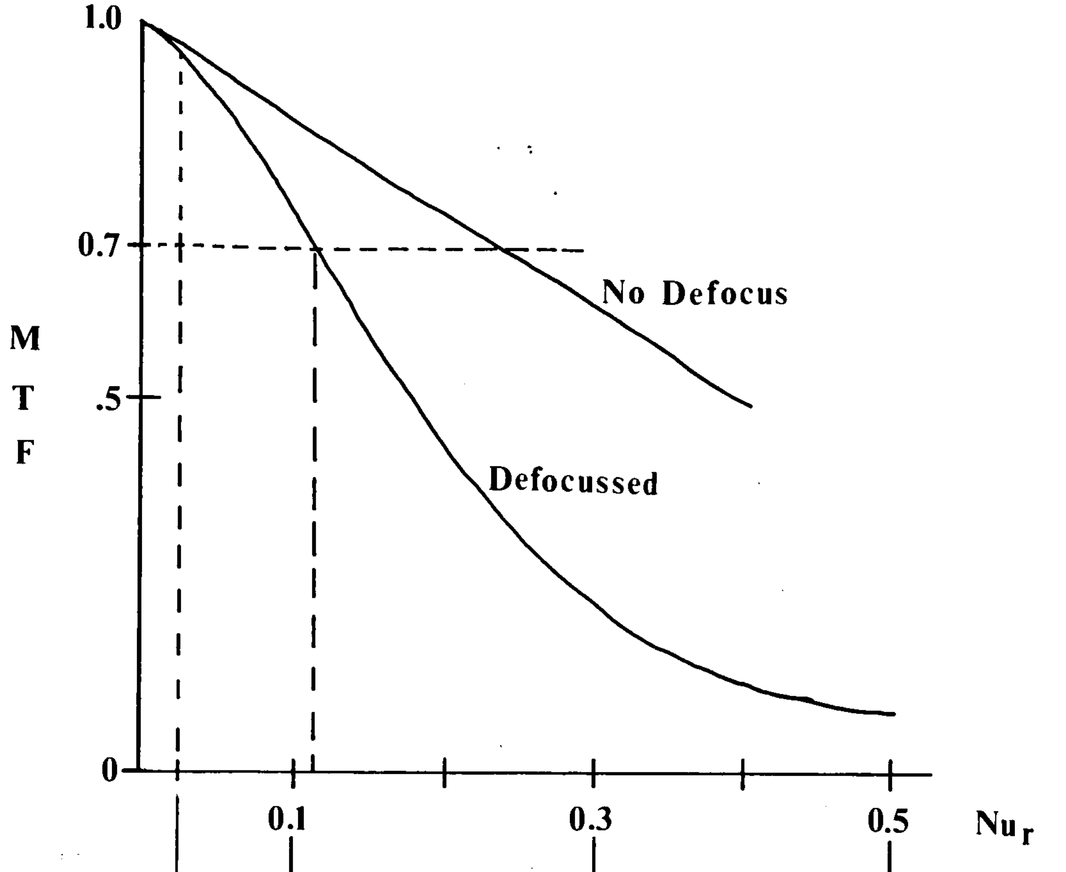
BC = Blur Circle

The hyperfocal distance was computed to be 5 meters and the near limit was 2.5 meters using a 0.1mm blur circle diameter.

MODULATION TRANSFER FUNCTION. Using Granger's subjective quality function, it was determined that the diffraction effects of the f-45 aperture would not be noticeable in contact prints.[19] Granger's subjective quality function states that if the modulation transfer function value (MTF) at 20 cycles/mm on the retina is 70% or higher then the viewer will not notice any degradation.

Because the eye operates at a magnification of approximately 1/20, a spatial frequency of 1 cycle/mm on the print will produce approximately 20 cycles/mm on the retina. A table of the MTF for a defocused, diffraction limited lens showed that even with 4.5 mm defocus that a MTF value greater than 95% was obtained at 1 cycle/mm.[20] Figure 3 shows that the 70% value was not reached until a spatial frequency of more than 4 cycles/mm.

The system MTF is the product of the MTF's of each separate part of the system: the optics, the film and its development, the printing, the paper and its development. Due to adjacency effects, the published MTF of Kodak Plus-X film is greater than one at a spatial frequency of 1 cycle/mm. The vigorous agitation used in this experiment, however, probably reduces this to nearly unity. The MTF of the Kodak Azo paper used is not published by Kodak, but it is assumed that the contributions of all other sources of MTF degradation are not sufficient to depress the system MTF below 0.70 at a spatial frequency of 1 cycle/mm.



F#			Lines/mm		Defocus
5.6	8	32	96	161	0.06
11	4	16	49	82	0.26
22	2	8	25	41	1.08
45	1	4	12	20	4.5

FIGURE 3

MODULATION TRANSFER FUNCTION FOR DEFOCUSED DIFFRACTION LIMITED LENS

This graph shows the MTF for a lens defocused by various amounts depending on the f-number. For a spatial frequency of 1 cycle/mm at f-45 and defocused 4.5 mm, an MTF of 0.95 is shown. The spatial frequency associated with an MTF of 0.70 under the same conditions is somewhat larger than 4 cycles/mm.

Airy Disk. The size of the Airy disk associated with the f-45 aperture

$$D = 2.44\ L\ f\# \qquad (4)$$

Where:

D = Diameter of Airy Disk

L = Wavelength of Light

f#= f-Number of lens

Using 555 nm for the wavelength, the diameter of the first zero of the Airy disk is 0.061 mm. This is roughly two thirds of the 0.1 mm blur circle used for the depth of field calculations.

Measuring the Aperture. The f-number of the lens was measured with the aid of a measuring microscope. The lengths of the edges of the polygon made by the iris edges were measured and calculating the radius of a circle of equal area was calculated using the following formulas:

$$A = (N\ L^2/4)\ \text{cotangent}\ (180/n) \qquad (5)$$

$$R = (A\ /\ pi)^{.5} \qquad (6)$$

$$f\# = FL\ /\ (2\ R) \qquad (7)$$

Where:

A = Area of Polygon

N = Number of Sides

L = Length of Side

R = Radius of Equivalent Circle

f#= f-Number of Lens

FL= Focal Length of Lens

The average of five determinations of f-number was 44.18.

MEASURING THE G-NUMBER. The G-number of a lens is defined as the ratio between the scene luminance and the film plane illuminance.[21] These quantities were measured using a UDT 40-x photometer. A special luminance probe was used that employs a well baffled aperture and provides a factor of 0.01 lux / meter lambert. The G-number found by this method was 9636. As the luminance and illuminance data had only 3 significant figures, the number of 9600 was taken for ease of computing

ALIGNMENT

The arrangement of three collimators mentioned above was used to align the camera. An 8x loupe was used to view the images of the Siemans stars on the ground glass. It was found that the zero detent position of the rear camera standard did not coincide with the best alignment. This part of the camera was replaced with a part that did coincide with the best alignment.

FOCUS

The focus was set by first focusing the Siemans star images to the best focus position. This effectively focused the camera at infinity. The film back of the camera was then moved towards the lens by an amount that left the camera focused at the hyperfocal distance of 5.000 meters. This was checked with test targets set at the hyperfocal distance.

LENS SHADE

A lens shade was constructed to minimize the amount of stray light in the camera and to provide for the changing of neutral density filters for the exposure control. A universal lens board was used that abutted the lens. A 3 inch square gelatin filter holder was mounted on the plate that supported the universal lens board. A telescoping bellows was used for the shade itself; the small end held by pyramidal stiffeners. The small end was masked with an aperture of the same aspect ratio as the usable area of the film. The position of the shade was adjusted carefully so that no part of it intruded into the picture at the taking aperture of f-45.

FILM PLANE MASK

A portion of the film plane had to be masked off to allow the printing of a sensitometer exposure on each piece of film. This was accomplished using a piece of tape that blocked a 10 mm wide strip across the top edge of the film (bottom edge of the picture). This provided for nearly square sensitometer exposures that were just large enough to be used by the densitometer.

PROCESSING

The processing requirements for this experiment were very stringent. The repeatability and uniformity had to be such that there was no detectable difference between two similarly processed negatives. Suggested processing methods included large batch, nitrogen burst, and tumble.[22] The uniformity between negatives in a large batch process did not seem guaranteed, and could take large amounts of materials to test. Nitrogen burst shared the uniformity problem. The tumble processor seemed the best solution even though it would process only one sheet at a time.

TUMBLE PROCESSOR

The tumble processor described to me had used a #1 olive bottle and an electric drill with a speed control. The bottle was just large enough to hold a strip of 35 mm film longitudinally and was filled one-third of its capacity with chemistry. The bottle was spun by the drill at a speed sufficient to cause vigorous, random agitation but not so fast to cause centrifugal force to keep the chemistry at the ends of the bottle. This process gave very uniform results and also was quite repeatable.

Obviously, to process 4x5 negatives, a container larger than a #1 olive bottle was necessary. A large peanut butter jar that was approximately 3.75 inches in diameter and 5.5 inches tall was chosen. This was attached to a mandrel with silicone rubber and was supported by a shaft with two ball bearings to reduce the radial loading on the motor. The large difference in size and shape between the peanut butter jar and the olive bottle necessitated a change in the rotational speed to obtain the necessary agitation. It was also found that the film would contact the side of the jar during processing. Some time was spent trying to devise a film holder to keep this from happening, but it was finally decided that the emulsion side was the side that needed the vigorous agitation during development and that the other side of the film could be cleared later if necessary.

UNIFORMITY AND REPEATABILITY

Film was uniformly flashed to produce densities around 2.5 when developed in D-76 for 7 minutes. The results showed that the differences between processing runs were insignificant compared to the uncertainty of the densitometer, but that the position of the density reading was significant. Looking at the film, it was possible to see a line that corresponded to the depth of the developer that was poured in and left standing while the lid was being put on. Another test was run and the developer was added while the jar was tilted to keep the negative from contacting the developer until the motor was turned on. This resulted in no effects that were significant compared to the uncertainty of the densitometry.

CHEMISTRY

Kodak's HC-110 developer was chosen because of the relatively short development times, and the ease of mixing different concentrations. Extremely dilute solutions would have risked oxidation due to the air in the processor and after complete development is acheived, higher concentrations merely raise the fog level. Times and concentrations used at 68° are shown in Table 1.

TABLE 1

KODAK HC-110 PROCESSING TIME AND CONCENTRATIONS

TIME	CONCENTRATION
9	1:15
9	1:31
11	1:63
3	1:63
3	1:128

At the conclusion of development, the developer was dumped out, stop bath put in and run for 15 seconds, followed by fixer for 3 minutes.

All negatives were printed on Kodak Azo paper using a point source and a contact printing frame. The exposed prints were processed in Kodak Dektol developer.

SENSITOMETRIC CONTROL

Sensitometric control during processing was provided several ways. Sensitometer exposures were added to each negative shortly after the camera exposure. All negatives receiving identical processing treatment were required to have no significant differences in transmission density on any step of the sensitometer exposures. Duplicate camera exposures were made to allow the replacement of any negative failing this test.

Each processing treatment also included one negative that had a reflection target inserted into the scene. This target was produced from flashed photographic paper. It was placed at the near limit of depth of field and the individual patches were sized appropriately for reflection densitometry. Reflection densitometry from this target were used to calculate the tone reproduction gradient for each print.

SCENE SELECTION

Three landscape scenes were chosen that fit the requirements of having only subtle visual clues. The scenes chosen were: Veteran's Bridge from the footpath in Rochester, N. Y. ; Mendon Ponds Park, Monroe County, N. Y.; and Hemlock Lake, Livingston County, N. Y. (see Appendix). The lighting of these scenes was frontlight, sidelight, and moderate backlight respectively. Visual clues available were stone, concrete, grass, water, and sky. In the case of the Mendon Ponds shot, shadows of trees were cast on the grass. And in the Veteran's Bridge shot, a large part of the bridge was in its own shadow. The Hemlock Lake shot had no significant shadows.

EXPOSURE

The exposure for each scene was determined by the luminance of a Kodak 18% Gray Card at the time of the exposure. The amount of neutral density needed to provide the illuminance necessary to produce a negative density of 0.74 was then computed from the gray card luminance and the lens G-number. Neutral density filters were then added or subtracted to provide a ± 0.2 range of relative log exposure. This range was chosen to accomodate the 0.1 resolution of the neutral density filters and the rather narrow range of exposure possible at the high reproduction gradients.

The exposures were made and sensitometer exposures added to each negative within 90 minutes. The processing began approximately two hours after the camera exposure. Contact printing of the negatives took place several days later. Sensitometric tests were made at each step for verification of control. Two scenes had to be reprinted because of an error in reading the tabular density information. It was later decided to include these incorrectly printed scenes in the judging.

MEASUREMENT AND SCALING

The concept of measurement is not easily defined despite the fact that everyone has a functional idea as to what it is. Bertrand Russell defined the:

> ...measurement of magnitudes ... in its most general sense (as) any method by which a unique and reciprocal correspondence is established between all or some of the magnitudes of a kind and all or some of the numbers, integral, rational, or real .[23]

Campbell defines measurement as:

> The assignment of numerals to represent properties of material systems other than numbers, in virtue of the laws governing these properties.[24]

Stevens merely states that:

> Measurement is the assignment of numerals to objects or events according to rules.[25]

The common thread that runs through these definitions is that numbers (or magnitudes) should be assigned to reflect an isomorphism between the number system used and particular attributes of the object measured. Three characteristics of the real number system are order, distance, and origin. Just as one may order various lengths of rope with the relation "larger than" the real numbers are ordered by the relation "larger than." The characteristics of distance may be thought of as the fact that differences between two numbers are ordered. Thus, the difference between numbers a distance 3 apart is larger than the difference between numbers a distance 2 apart. The number system has a unique origin at the number zero, the additive identity.

Typically, some isomorphism between these characteristics and the attribute of concern is recognized by an empirical meaning attached to the isomorphism, although this is not necessary. For example the lengths of rope mentioned earlier could be named with a number thus creating a nominal scale. If these names reflected the longer lengths with larger numbers, an ordinal scale would have been created. If it were important to know the actual lengths involved, the differences between all the lengths could also be ordered creating an interval scale of arbitrary unit. If we were to assign the number zero to ropes that have no length, a ratio scale results. A ratio scale would enable the ordering of all ratios of lengths of rope.

Besides the types of scales, the type of measurement must be decided. This decision determines the relationship through which the isomorphism gains its meaning. These relationships may: depend on fundamental or natural laws relating the various quantities, or be derived from laws that relate the quantities, or may be stated without proof.

The present state of psychophysics rules out the first possibility of fundamental measurement of psychophysical quantities because of the lack of fundamental or natural laws relating the quantities. All psychophysical measurement, then, requires that either a model be constructed that relates the various attributes to each other, or that a measurement be defined to represent a relationship without proof (by fiat).

Three classes of psychophysical measurements can be defined depending on whether the variation in the measurement is attributed solely to the subject, the stimuli, or both. The familiar classroom test assumes that differences between pupils is due to the variability in the pupils. The example of ordering the lengths of rope, on the other hand, assumes that differences are due to real differences in the stimuli. The third possibility and the most general is the possibility that if the rope scaling experiment were carried with several different sets of ropes, the variability in the measurements might be due not only to the subjects scaling the ropes, but might be due to the variabilities in the sets of ropes. It should be noted that when the variability of a general psychophysical measurement is attributed solely to the subjects, that no widely applicable scaling methods are available. Measurements of this type reduce to measurement by fiat. In the stimulus centered approach, the subjects are asked to judge the stimulus in relation to other stimuli within the psychological continuum considered in an attempt to minimize bias. The response approach, as used in this experiment, merely asks the subject to judge the stimulus in relation to his own attitudes regarding the psychological continuum.

LAW OF CATEGORICAL JUDGEMENT

Torgerson states that the law of categorical judgement is based on Thurstone's general judgement model.[26] That is, there is a postulated psychological continuum, that a stimulus causes some sort of discriminal process that has a value on the continuum, that the stimulus is not always associated with one value on the continuum, that these values possess a normal distribution, and that different stimuli may have different means and different standard deviations. With the addition of one assumption, the law of comparative judgements is derived from the general judgement model: When 2 stimuli are presented to a subject, he will judge the first higher than the other when it has a higher value on the continuum.

To extend the law of comparitive judgement to the judgement of categories, this last assumption must be changed to:

1. The psychological continuum of the subject can be divided into a specified number of ordered categories.
2. A given category boundary is normally distributed and different boundaries may have different means and standard deviations.

3. The subject judges a stimulus to be below a given boundary when the value of the stimulus is below the boundary on the continuum.

This amounts the assumption that boundaries behave like stimuli.

The law of categorical judgement may be stated mathematically:

$$t_g - S_j = x_{jg} \left(d_j^2 + d_g^2 - 2r_{jg}\, d_j\, d_g \right)^{.5} \qquad (8)$$

Where:

t_g = mean location of g^{th} category boundary

S_j = mean location of j^{th} stimulus

x_{jg} = unit normal deviate corresponding to proportion of times s_j is sorted below boundary g

d_j = standard deviation of j^{th} stimulus

dg = standard deviation of g^{th} boundary

r_{jg} = correlation between momentary position of stimulus j and category boundary g

In complete form these equations are not soluble since there are more unknowns than equations. With n stimuli and m+1 categories, there are 2 (n+m-1) + m n unknowns and only m n equations. In order to simplify the problem, certain of the terms may be assumed constant. One of the least objectionable assumptions is that the standard deviations of the stimuli are equal as this has disasterous results only surrounding the rather uninteresting proportions of 0 and 1. Computer programs are available that will solve the resulting equations iteratively.

The experimental procedures that are acceptable with this method involve either sorting the stimuli into m + 1 piles where the piles are in rank order, or rating a stimulus one at a time using numeric, adjective, or graphical scales, or a straight ranking of the stimuli in order. Of the three alternatives, the rating is the easiest to perform and requires the least skill from the judge.

SCALING OF INDIVIDUAL DIFFERENCES

Many disciplines are concerned with the psychophysical response of people to various stimuli. Because of the thousands of possible factors involved in having people judge a psychophysical quantity directly, most tests have relied on judging the difference between two stimuli that are only slightly different. This approach has led to the concept of the "just noticeable difference" used in the fields of color, acoustics and many others.

This type of analysis requires a knowledge of the dimension being investigated, that is ,that the given stimuli do in fact vary in a specific manner within the dimension concerned. Multidimensional scaling, in general, requires no specific knowledge of the underlying dimension(s) involved. Given only data describing the similarity or dissimilarity of each pair of stimuli that may vary in an unspecified manner, a multidimensional scaling of the data will determine the dimensionality of the underlying structure.

One of the most frequently used examples of this is to use the matrix of intercity distances from a map as a measure of similarity or proximity.[27] The use of any one of several multidimensional algorithms on this data usually produces an array of points representing cities on a two dimensional "psychological continuum" that looks remarkable like the original map. The map has been reconstructed from the intercity distances. Likewise, the intercity distances may be reconstructed very accurately using the positions of the points on this continuum, a much smaller set of data.

The mechanism of reducing the data is not conceptually difficult to understand. While there are several slightly different approaches to the solution, most share the approach put forward by Shepard or a modification of it. Shepard's methodology followed three steps:[28,29]

1. Adjust the lengths of the sides of a regular N-1 dimensional figure until the inter-vertex distances bear an monotonic relationship to the inverse rank order of the original similarity data. This can always be done in N-1 dimensions.

2. Treat each inter-vertex distance as if it were a force vector and move each vertex along the resultant vector a distance that is proportional to the mean distance.

3. Rotate the coordinate axes so as to eliminate the projection of the data in at least one dimension.

The first step guarantees a monotonic relationship between the raw data and the vertices of the figure being postulated. The second step tends to diminish the dimensionality of the figure by forcing vertices that are close even closer to each other and those that are far away even farther from each other. The third step merely aligns the coordinate axes with the major trends of the vertices and actually expresses the reduced dimensionality in fewer dimensions.

The INDSCAL procedure used in this experiment, however, minimizes a multi-dimensional stress vector by least squares iteration.[30,31] This stress vector represents the mean square departure of the data from the hypothesis that the measured interpoint distances differ from the true interpoint distances only because of random fluctuation. Because INDSCAL does not optimize the dimensionality of the solution, it is left to the experimenter to determine the optimum dimensionality of several solutions provided by the program.

JUDGING OF STIMULI

The stimuli were judged by 10 volunteers in convenient surroundings. No attempt was made to provide specific conditions, but information on each condition was taken for possible later correllation with the data. As mentioned previously, some judges had a sheet with seven adjectives marked on it as below:

1. Unusable
2. Poor
3. Unsatisfactory
4. Acceptable
5. Good
6. Very Good
7. Excellent

These adjectives had been used in similar experiments before and have been found useful. The judges that were not given the list were asked to rate each picture from 1 to 7 with 7 being the best and 1 being the worst.

The small number of judges was thought adequate based on Granger's experience of comparing the results of small subgroups of judges with the total group of judges. The amount of data that is generated for the INDSCAL analysis from even so small a sampling of the population is quite large. Each scene produced 300 measures of pair proximity for each subject, or 1,500 measures for the 5 scenes. This produced 15,000 measures of pair proximity for the experiment. The initial iteration of the algorithm assumes a 299-dimension space for each scene and a correspondingly large number of equations. The volume of the calculations involved in this type of analysis preclude the analysis of truly large samples of the population. Because INDSCAL displays the relative importance that each subject gave to each dimension, it is possible to detect subjects that "marched to the beat of a different drum."

The prints were held in a stack by the subjects and were presented in a random order. The data taken in addition to the rating of each print were the illumination, whether the room "felt" light or dark, whether a key white was available to the subject during the judging, whether a list of adjectives was available to the judge, and whether the subject seemed nervous or not.

RESULTS

The categorical data from the judges were used as input for a program that determined the boundaries and the variance of the boundaries of the categories on a psychological continuum. These data are presented graphically in Figure 4. The category boundaries are very nearly 3 times the standard deviation of the boundary positions giving categories that are significantly different at the 95% confidence level. Most are also significant at the 99% level. The average ratings for each print were plotted against the 5 levels of exposure and contrast. Using the values for the boundaries found above, a contour plot was drawn (Figure 5). Plots were also made of average quality as a function of exposure and average quality as a function of contrast level (Figures 6 and 7).

Figure 6 showed that there was no significant difference in average quality between the two highest contrast levels. Checking of the tone reproduction curves confirmed that there was no significant difference in contrast between these levels. Therefore, contrast levels 4 and 5 have been grouped together.

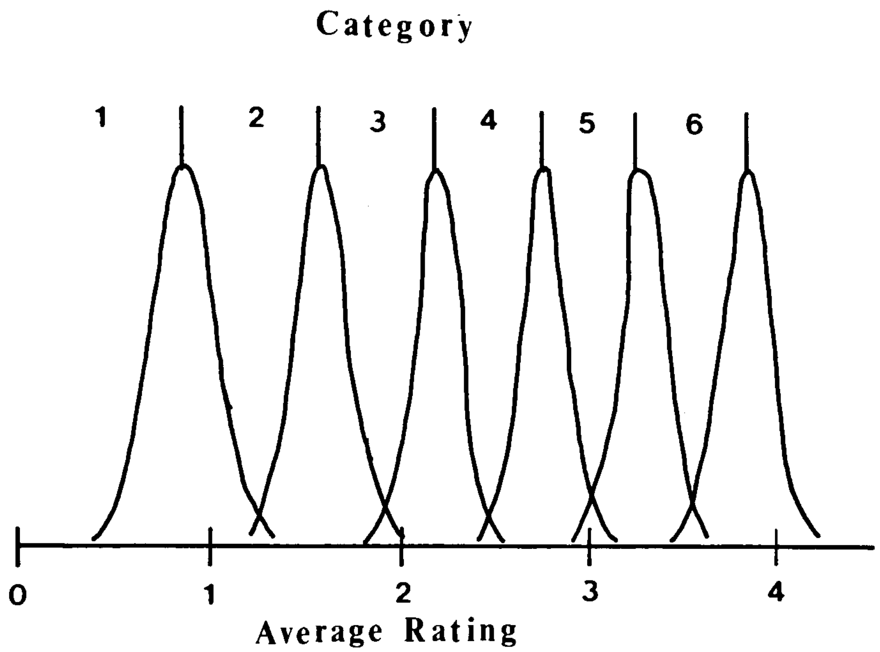


FIGURE 4. DISPERSION OF CATEGORY BOUNDARIES

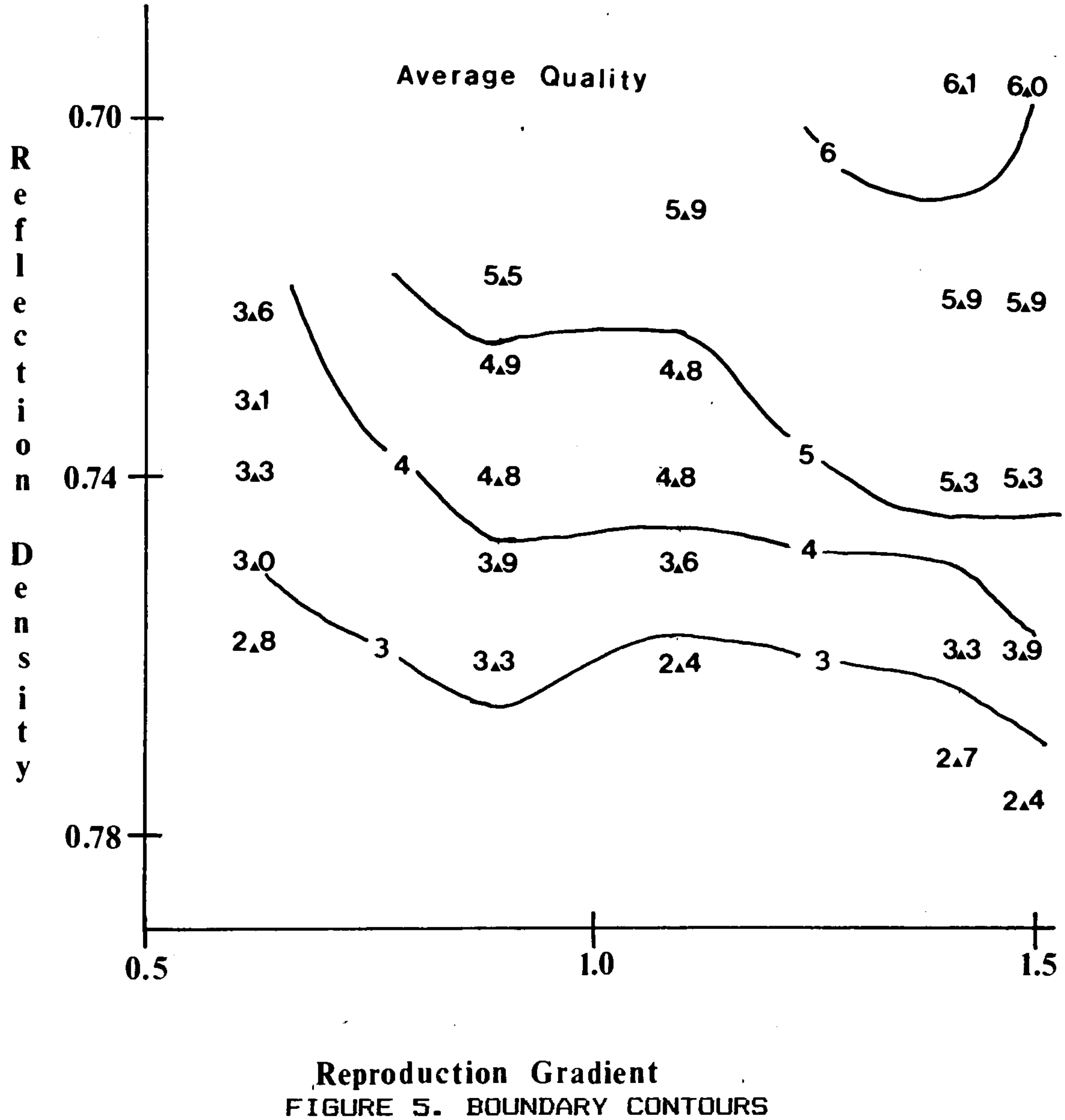


FIGURE 5. BOUNDARY CONTOURS
Average quality contours are shown in relation to the reflection density of an 18% gray target and reproduction gradient.

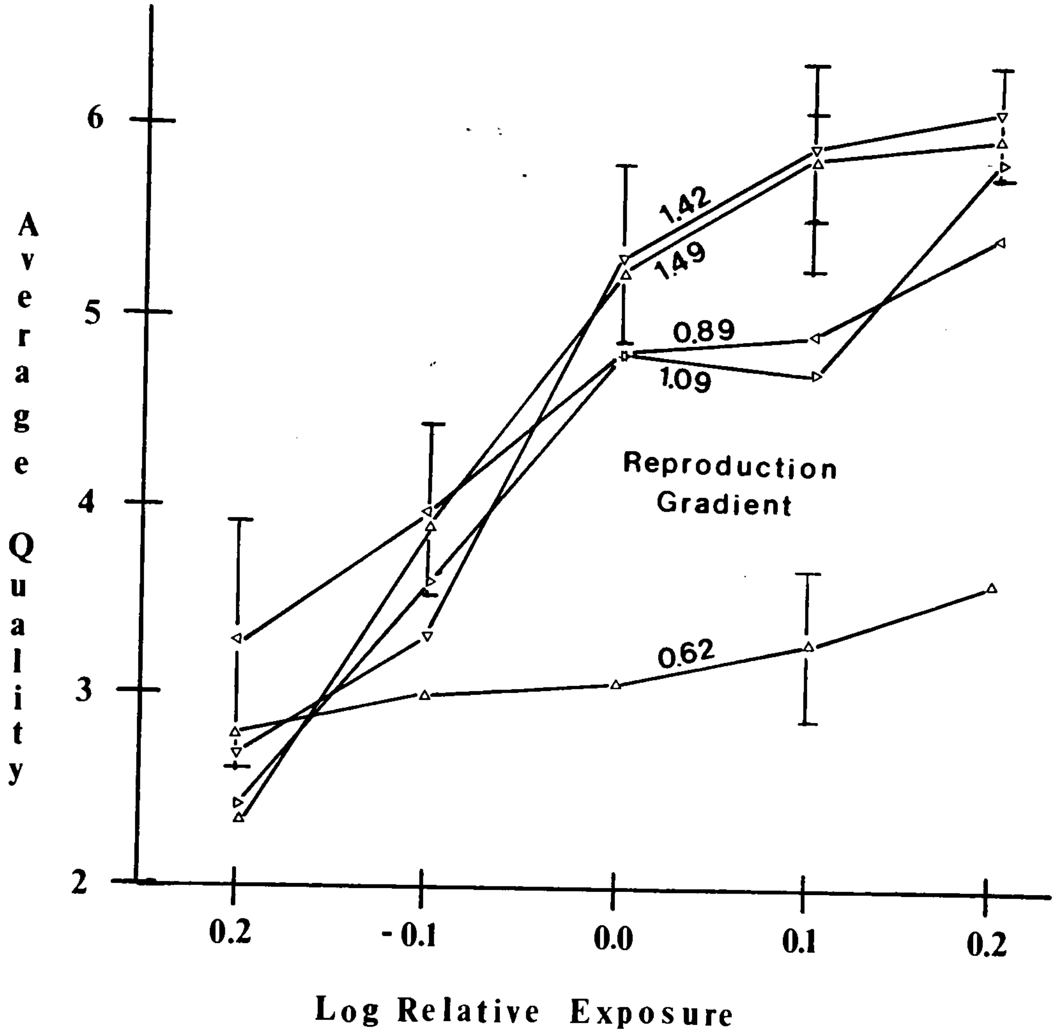


FIGURE 6. AVERAGE QUALITY vs: EXPOSURE LEVEL
The average quality as a function of negative exposure is shown for each reproduction gradient.

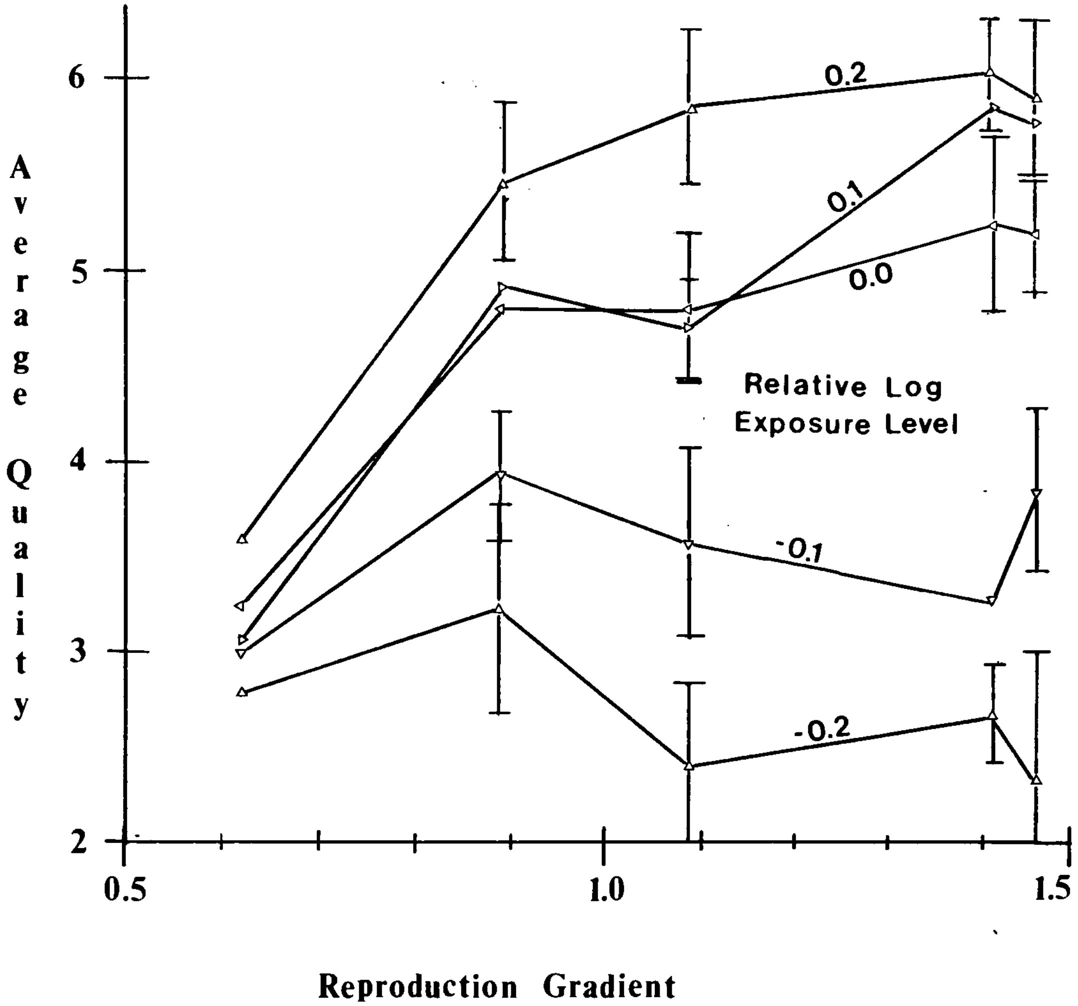


FIGURE 7. AVERAGE QUALITY vs: CONTRAST LEVEL
The average quality as a function of reproduction gradient is shown for each exposure level.

The average tone reproduction curve gradients are listed in Table 2.

TABLE 2
AVERAGE TONE REPRODUCTION CURVE GRADIENT

LEVEL	GRADIENT	90% CONFIDENCE INTERVAL
1	0.620	±0.298
2	0.886	±0.075
3	1.090	±0.136
4	1.416	±0.154
5	1.490	±0.349
4&5	1.456	±0.245

Similarly, the plot of average quality as a function of exposure showed that at most contrast levels there was no significant difference between exposure levels 4 and 5. The relative exposures and the associated effect on the density of an 18% gray target is shown in Table 3.

TABLE 3

DENSITY CHANGE OF 18% GRAY DUE TO EXPOSURE LEVEL

REL EXP LEVEL	LOG EXP	CONTRAST LEVELS 1	2	3	4	5
1	−0.2	+0.18	+0.22	+0.30	+0.44	+0.44
2	−0.1	+0.08	+0.12	+0.12	+0.20	+0.20
3	0.0	0	0	0	0	0
4	+0.1	−0.10	−0.10	−0.15	−0.19	−0.19
5	+0.2	−0.19	−0.21	−0.26	−0.31	−0.36

A measure of pair proximity was produced from the categorical data by subtracting the category that each print was ranked from the category of each other print in that scene. The difference was a measure of how dissimilar the prints were. The dissimilarity data were used as input to INDSCAL, an individual differences scaling program. Several runs were made with different numbers of dimensions in the solution. These results were disappointing. The solutions of only a few dimensions did not explain very much of the variability in the data (see Table 4) and the solutions with many dimensions were nearly impossible to interpret.

TABLE 4

PERCENTAGE OF VARIANCE EXPLAINED BY INDSCAL

NUMBER OF DIMENSIONS	VARIANCE EXPLAINED
2	68.07
3	72.39
4	74.47
5	76.98

The two dimensional solution accounted for only 68% of the variance of the data and a five dimensional solution accounted for only 77%. Moreover, the incremental variance accounted for by each successive dimension was relatively constant suggesting that more dimensions than five may be necessary to fit the data. The plots of the stimuli and subjects in the many-dimensional space that INDSCAL provided were used to determine if any judges differred significantly. There seemed to be no correlation between any of the supplemental data taken at the time of judging and the positions of the judges in these spaces. In several dimensions the judges who were known to be experienced observers of photographs were positioned close to many of the least experienced and most nervous judges.

CONCLUSIONS

In each scene, the subjects preferred the most contrasty prints offered to them. However, in one scene due to an error in printing exposure, a gradient of only 1.16 was produced for the most contrasty prints. This gradient has been dropped out of the averages reported on the basis that the maximum contrast presented was too low to test the hypothesis that the preferred gradient had not changed.

The average contrast of the most preferred prints was 1.456 ±0.245 at the 90% confidence level. Previous studies using black and white prints found a midpoint gradient of 1.10 and an average gradient of 0.9 . This figure is an average of the preferred tone reproduction characteristics from Jones and Nelson and Jones and Condit as reported by C. N. Nelson.[4] In the third edition of the same work, Nelson reported that this average ranged from 1.10 to 1.20.[9] In either case, the average found in this study is significantly different at the 90% confidence level.

The most preferred prints were from negatives that received approximately 1.6 times the exposure that is necessary to reproduce an 18% reflectance target as 18% reflectance. That is, the subjects most preferred middletone densities that were 0.30 to 0.36 lighter. This is consistent with previous findings.

The preferred mid-point tone reproduction gradient found in this study (1.46 ±0.24) is different at the 90% confidence level from that found to be preferred in previous studies (1.10-1.20). The hypothesis that the preferred tone reproduction gradient to be found in this experiment is equal to the preferred tone reproduction gradient found previously is disproved and the thesis is supported. The amount of exposure required for the most preferred prints in this study did not significantly deviate from previous studies.

The application of the results of this experiments to more general situations is complicated by several factors. The exclusion of obvious visual clues of tone values such as skin or white clouds is an unusual way to take a photograph. Very few photographs are contact printed because of the popularity of small camera formats. However, most prints are held in the hand and viewed under whatever illumination is available.

The small number of subjects and scenes used in this study is a major concern in the application of these results to a more general situation. In most cases, to extrapolate to a population of millions from a sample of only ten is foolhardy. The science of psychophysics is by no means exact. Many of laws of psychophysical measurement are controversial. The INDSCAL analysis of the data failed to be useful.

However, it takes only the existence of one black sheep to disprove the hypothesis that all sheep are white. There is evidence in this experiment that shows that under the conditions outlined, ten subjects did prefer a print with a higher contrast than earlier studies indicated. The results of this experiment are consistent with independent work by W. R. Dowling using several scenes that included visual clues and used projection prints made from 35mm negatives.[32] The results are also consistent with the findings that prompted this study in the first place.[8] Based on these facts, it is my feeling that these results are applicable outside the rather narrow confines of this work.

RECOMMENDATIONS FOR FURTHER WORK

One of the most obvious areas to improve on would be to perform the experiment in an iterative or sequential fashion. This would allow the area of exposure-contrast space that would be of greatest interest to be identified before the massive production of the final viewing stimuli. It would also allow timely feedback for eliminating camera flare.

Producing the tone reproduction curves from actual scene luminances rather than the use of a reflectance chart inserted into a scene would be a large step forward. Unless the measurement of the reflectance factor of the chart duplicates the geometry of the lighting and pickup by the camera of the chart when it is used in the field, errors may be introduced that may be large depending on the surface characteristics of the chart.

It is my belief that the INDSCAL analysis failed because of the imprecision involved in artificially creating the paired comparison data. From the categorical analysis of the data it can be seen that the standard deviations of the positions of the boundaries are fairly large with respect to the distance between boundaries. The generation of difference data, in retrospect, seems only to make matters worse. To begin with, the data were integers between 1 and 7. A true paired comparison would resolve differences between stimuli that are lost in the poor resolution of the present method. However, the number of stimuli would have to be reduced to have the comparisons take place within a reasonable amount of time.

The number of scenes should be increased. The results from each scene and could be weighted based on recent studies of "system utilization space" that have characterized the frequency of pictures taken as a function of scene luminance and subject distance.[33] It should be noted that the scenes used in this study fall in the vicinity of the highest frequency of pictures taken, that is, a nearly infinite subject distance and full daylight illumination.

The number of subjects should also be increased. The exact number will depend on the amount of variability built in to the experiment by the stimuli used and the variability that is tolerable in the final results. This quantity may be determined from the data that would be available if the experiment is done in a sequential fashion.

REFERENCES

1. E. P. Wightman, "Photographic Exposure Guides and Meters," Photographic Science and Engineering, 2(1958):14-31

2. R. Taft, Photography and the American Scene, (New York: Macmillan, 1938; Dover Publications, 1964)

3. W. B Ferguson, ed., The Photographic Researches of Ferdinand Hurter and Vero C. Driffield, (New York: Morgan and Morgan, 1974)

4. C. N. Nelson,"Tone Reproduction," in The Theory of the Photographic Process, 4th ed., edited by T. H. James, (New York: Macmillan, 1977) 536-560

5. L. A. Jones and C. N. Nelson, "The Control of Photographic Printing by Measured Characteristics of`the Negative," J. Opt. Soc. Am., 38 (1942): 558-619

6. L. A. Jones and H. R. Condit, "Sunlight and Skylight as Determinants of Photographic Exposure I. Luminous Density As Determined by Solar Altitude and Atmospheric Conditions," J. Opt. Soc. Am., 38 (1948):123-178

7. L. A. Jones and H. R. Condit, "Sunlight and Skylight as Determinants of Photographic Exposure II. Scene Structure, Directional Index, Photographic Efficiency of Daylight, Safety Factors, and Evaluation of Camera Exposure," J. Opt. Soc. Am. ,39 (1949):94-135

8. E. M. Granger, Private Communication

9. C. N. Nelson, "Tone Reproduction," in The Theory of the Photographic Process, 3rd ed., ed. C. E. K. Mees and T. H. James, (New York: Macmillan 1966) 467

10. G. T. Fechner, Element der Psycophysik, vol I,II, (Leipzig: Breitkopf and Hartel) 1889

11. R. B. Marimont, "Model for Visual Response to Contrast," J. Opt. Soc. Am.,52 (1962):800-806

12. J. L. Simonds, "Analysis of the Variability Among Density-Log exposure Curves of Black and White Negative Films by the Method of Principal Components," Photogr. Sci. Eng. , 2 (1958):205-209

13. L. D. Clark, "Mathematical Prediction of Photographic Picture Quality from Tone Reproduction Data," Photogr. Sci. Eng. , 11 (1967):306-315

14. C. N. Nelson, "Safety Factors in Camera Exposures," Photogr. Sci. Eng., 4 (1960):48-58

15 A. Adams, The Negative Exposure and Development, New York: Morgan and Morgan, 1968

16 M. White, Zone System Manual, New York: Morgan and Morgan, 1968

17. J. L. Tupper, "General Sensitometry," in The Theory of the Photographic Process, 3rd ed., ed., C. E. K. Mees and T. H. James,(New York: Macmillan):1966

18. W. J. Smith, Modern Optical Engineering, (New York: McGraw-Hill, 1966)

19. J. C. Carson, Private Communication

20. L. Levi and R. H. Austing, "Tables of the Modulation Transfer Function of a Defocused Perfect Lens," Applied Optics, 7(1968):967-974

21. A. Stimson, "The G-Number: A Photometric Lens-Aperture designation," J. Soc. Motion Picture Television Engrs., 74(1965):99-101

22. R. Francis, Private Communication

23. B. Russell, Principles of Mathematics, (New York: Norton, 1938)
24. Proceedings of the Aristocratic Society, Suppl. 17(1938):121-142

25. S. S. Stephens, "Mathematics, Measurement and Psychophysics" in Handbook of Experimental Psychology, ed., S. S. Stephens, (New York: Wiley, 1951)

26. W. Torgerson, Theory and Methods of Scaling, New York: Wiley, 1958

27. F. W. Young, Alscal-4 Collected Papers, (Carrboro,N. C. : Data Analysis and Theory Associates, 1979)

28. R. N. Shepard, "The Analysis of Proximities: Multidimensional Scaling with an Unknown Distance Function I.," Psychometrika, 27 (1962):125-140

29. R. N. Shepard, "The Analysis of Proximities: Multidimensional Scaling with an Unknown Distance Function II.," Psychometrika, 27 (1962):219-246

30. J. B Kruskal, "Nonmetric Multidemensional Scaling: A Numerical Method," Psychometrika, 29 (1964):115-129

31. L. R. Tucker and S. Messick, "An Individual Differences Model for Multidimensional Scaling," Psychometrika, 28 (1963):333-367

32. W. R. Dowling, "Derivation of the Optimum Film Contrast Gradient in Photographic Tone-Reproduction Systems," (B. S. Thesis, Rochester Institute of Technology, 1982)

33. T. M. Rice and T. W. Faulkner, "The Use of Photographic Space in the Development of the Disc Photographic System," Journal of Applied Photographic Engineering, 9(1983):52-57

APPENDIX

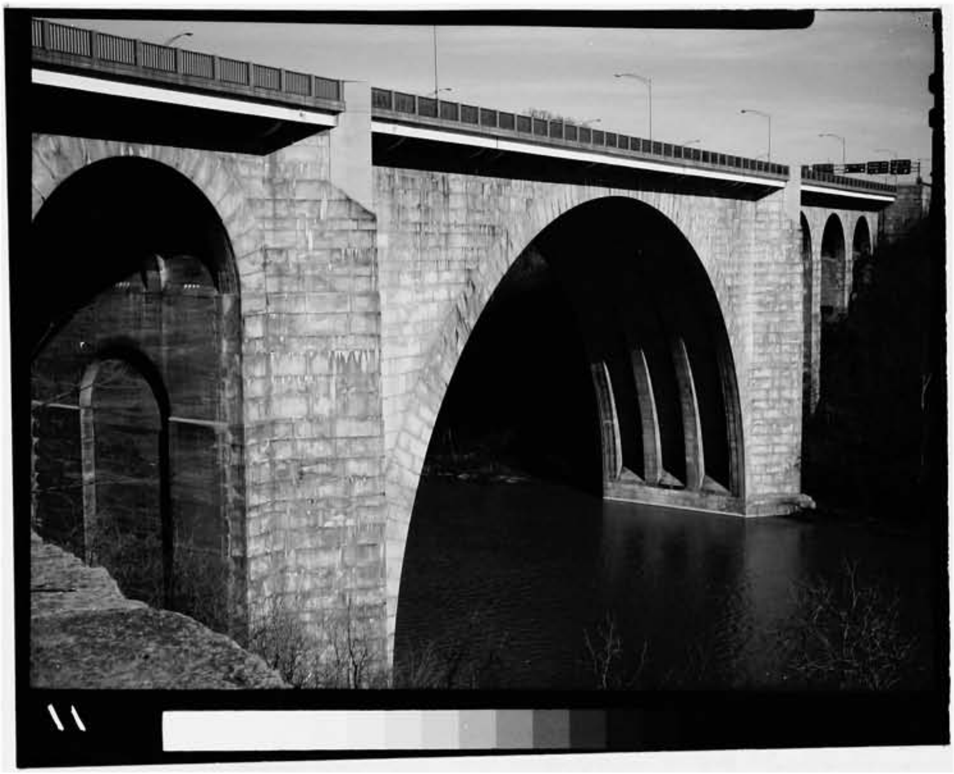


FIGURE 8.

Veteran's Bridge, Rochester, New York

FIGURE 9.

Mendon Ponds Park, Monroe County, New York

FIGURE 10.

Hemlock Lake, Livingston County, New York

VITA

David J. Porter is presently a United States citizen and was born in Coral Gables, Florida on March 29, 1947. He received a Bachelor of Science degree in the Philosophy of Science from North Carolina State University in 1972. Presently, he is employed in the Basic Standards Laboratory of the Macbeth Division of Kollmorgen Corporation.